Proceedings of the Annual General Donald R. Keith Memorial Conference
West Point, New York, USA
May 4, 2023
A Regional Conference of the Society for Industrial and Systems Engineering

Escape Room Efficiency and Integration of Automation

Emily Bancroft, Jared Culberston, Cara Dennehy, Omar Maflahi, Jack Martinson, Corey Ryan, and Daryl Santos

Department of Systems Science and Industrial Engineering
Binghamton University
Binghamton, New York 13902

Corresponding Author's Email: ebancro1@binghamton.edu

Author Note: Emily Bancroft, Jared Culberston, Cara Dennehy, Omar Maflahi, Jack Martinson and Corey Ryan are students at Binghamton University Watson College of Engineering and Applied Science. The contents of this report were made possible by support by capstone advisor Dr. Daryl Santos of Binghamton University's Department of System Science and Industrial Engineering.

Abstract: Escape rooms have been growing in popularity in recent years as an activity where groups work together to solve puzzles and escape a room. To monitor these games, a Game Master is required to watch the game through video surveillance to see what help the group needs. A local escape room center with four games and two control rooms is hoping to expand and hire new Game Masters coming in with no experience. This effort utilizes a suite of industrial and systems engineering techniques including, but not limited to time studies, discrete event simulation, and systems monitoring to enable the client to predict the progress of groups as they make their way through the games. As a result, both experienced and inexperienced Game Masters will be able to better monitor their rooms and administer hints while still giving groups a chance to escape.

Keywords: Escape rooms; time studies; simulation

1. Introduction

The project work is being performed for an escape room center that runs as a fundraiser for a 501c-3 non-profit that supports people with disabilities. All proceeds that are generated from the escape room center go towards the non-profit. Teams of 4-8 individuals work together using clues to solve a series of puzzles and escape the themed room. With the addition of a new room this year, five rooms will be run out of two control rooms, with one Game Master (GM) that can monitor up to two rooms at a time. The monitoring of teams from the control room entails watching up to 16 cameras at once for each themed game. As a team collects hints, they may finish one phase of the escape room and then move into another room, but may have to return to a previous room for more clues later.

The center also plans to hire volunteers that will receive community service time in exchange for helping, specifically with performing the role of a GM. This will mean that there will be a flow of new workers with low experience. The project objective is to monitor the current process to predict times to send hints to teams in the room, in order to make monitoring easier for a GM at any experience level. The focus for the simulation modeling being performed is on the most popular game at the center.

2. Methodology and Literature Review

The research team worked with the client to understand current restrictions to the way games are monitored to decide how best to apply engineering methodologies. It was decided that system monitoring and simulation would be the most appropriate way to model the current process. Discussion with the client showed that there was not a good system in place to tell how far along a group should be in the game, and that currently there was no log of hints that are given frequently in games. Because of this, it took a few game run throughs in order for a new GM to be trained, as the games require significant assistance in order to progress and eventually escape. To get better data on the rooms and progression of teams, it was decided that time studies should be performed with the eventual goal of simulating the times it should take to finish each puzzle to escape the room in a reasonable time. The center did not have many games running per week during the data collection period, so the goal was to collect as many data points as possible. Using too few points in a time study increases the chance of Type II errors, or accepting the null hypothesis although it may not be true (Faber & Fonesca, 2014). To avoid this, the team recorded as much data as possible and took notes on any outside factors that may have affected the data creating outliers. Using this data, the team can create a manual for future GMs that will allow for a simpler training period for people of all experience levels. The simulation will show what puzzles are taking too much time and may need to be modified in order for a better customer experience.

Spending time in the control room performing time studies allowed for the opportunity to collect data on what hints are given, and what time they are given at. The hint data that is gathered will then be compiled into prerecorded audio hints that can be used by the GM and are themed to the room being played. Studies on escape rooms in education have shown that not interrupting the immersion into the game can help young minds work better to solve the puzzles (Veldkamp et al., 2020). In order to not break the immersion in the game, the hints will use terminology and voice that fit the theme of the room. The partial shift to prerecorded audio should not affect the players' performances, as it has been found to be just as effective as live discussion in other fields. For example, a 2018 study on CPR performance showed that "the chest compression-only CPR guided by the pre-recorded instructional audio is no less efficient than dispatcher assisted CPR" (Birkun et al., 2018) and while it was in a different setting, the same principles apply to live hints versus pre-recorded hints being relayed from the GM to players in real time. Live GM assistance will still be necessary due to the personable and unique nature of escape rooms.

3. Results and Analysis

3.1 Data Collection

In order to collect data in a time trial format, the research team observed games as they took place, making sure to have at least one team member at as many games as possible. They watched as the GM ran the game and recorded the players' times at certain checkpoints. These checkpoints were provided to the research team by the GM, as the client had previously conducted a limited number of their own time trials with predetermined checkpoints. These checkpoints are typically each individual puzzle within each game, which were also broken up by each section within each game. An overhead view of the most popular escape room appears as Figure 1.

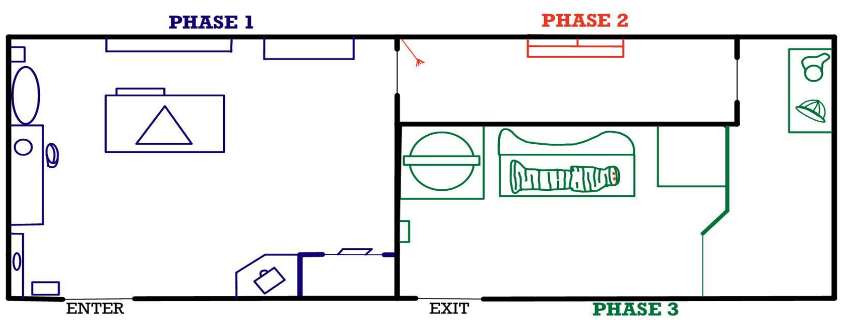


Figure 1: Layout of Most Popular Game: Teams work together through the three phases, using clues around the room to progress. Cameras monitor guests progress, but it can be initially difficult to recognize what puzzles are complete as a GM.

These checkpoints typically ranged anywhere from four to 10 per section and approximately twenty per game, providing the research team with a sufficient amount of data to analyze. The research team then recorded in a spreadsheet how much time was remaining in the game as the group of players reached each checkpoint, counting down from either 60 or 90 minute depending on the game. From here, the research team then found the means and task time for each puzzle or set of puzzles. Sets of puzzles were used when certain puzzles could be completed without a specific order, which differs from the linear nature of most other puzzles.

Another challenge when collecting data was that a majority of teams do not complete their game. The research team's goal was to record at least 10 attempts per room in order to receive sufficient data, however not all groups finished all of the puzzles in each room. This is not a problem for the client since their games are designed to not be escaped 100% of the time, with some games' success rate being less than 50%, but did mean that puzzles later in the game had less data than those at the start.

3.2 Data Filtering

For the purpose of applying data that was truly relevant to determining the ideal pace of puzzle completion within an escape room, it is important to interpret the limitations of the data that the research team had collected. Only being able to

ISBN: 97819384964-4-8

collect data from at most ten games per room due to restrictions and the majority of the teams that were observed failed to successfully escape their room, thus all the data that the team collected from these groups were not representative of an ideal pace. Because the goal is not necessarily to get the groups to escape as quickly as possible, but rather to have groups experience a scenario where they can experience the most competitive game, the team also had to consider the misleading data that is derived from teams that finished well before the allotted time. This prompted the team to utilize data filtering techniques that consider the mean and standard deviation of the data to evaluate outliers case by case (Winkler, 1993). In an ideal setting, if the research team had the ability to collect data from an unlimited number of games, the team would only consider groups that were close to finishing or finished with little time to spare.

Because of a limited dataset, the research team was forced to manipulate the data that was available. To do this, the research team first found the average time spent solving each puzzle, discarding data points that were more than three standard deviations away from the mean (1). The research team then took these averages and added them together to determine the observed pace in which an average team completes a game (2). Because most teams that were observed failed to escape the rooms, the average puzzle times were greater than the allotted time that was given for a team to complete the game. For example, for the most popular room, the summation of average time to complete all puzzles was 66 minutes, which was greater than the time given for the game the research team focused on, which is an hour (3). To account for this, the team divided the total puzzle averages by the amount of time given per game. Our scale ratio came to be .899, meaning that in order for the average team to complete the game, they must spend 89.9% or less of the time that the observed teams did (4). This ratio was then multiplied to each puzzle average so that when added together, would now add to a value that was equal to the time allotted per game, which in this case equals 60 minutes (5).

$$Average\ Observed\ Puzzle\ Completion = \frac{\sum_{i=1}^{n} Observed\ Puzzle\ Completion}{n} \quad (1)$$

$$Observed\ Pace = \sum_{i=1}^{n} Average\ Observed\ Puzzle\ Completion = 66\ min \quad (2)$$

$$Ideal\ Pace = 60\ min \quad (3)$$

$$\frac{Observed\ Pace}{Ideal\ Pace} = Scale\ Ratio = .899 \quad (4)$$

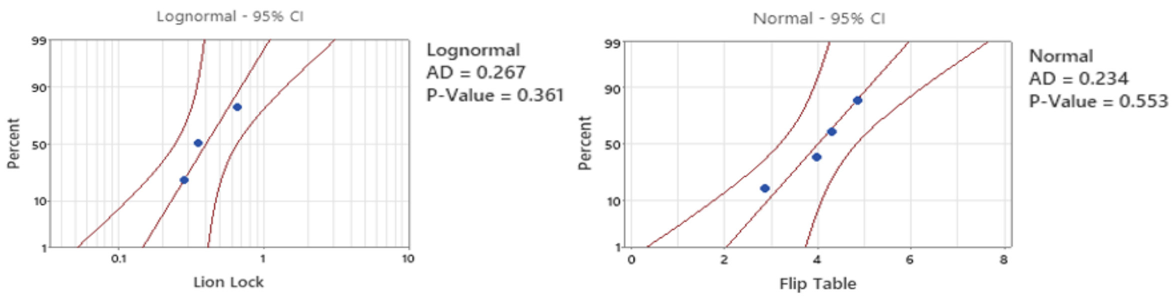
$$Ideal\ Puzzle\ Completion = Observed\ Puzzle\ Completion * Scale\ Ratio \quad (5)$$

One thing to keep in mind because of the data that was presented is the weighted difference between puzzles completed early in a game vs. the puzzles completed later in a game. Because the game ends at the expiration of the timer, there are some puzzles that a team may never reach. This is important when considering the data because the first few puzzles give a good representation of any random group of teams, while the last few puzzles in a game were only attempted by teams who had completed the game, thus typically being teams that can be expected to complete puzzles quicker. In order to combat this, the research team used heuristic measures to give slightly less time in the beginning puzzles for the suggested pace and more time for those at the end. A study done on a escape room performed in 2019 stated that "after the player goes through the learning period, which is about one fourth (1/4) of the total playing time, it can be seen a drastic decrease in the level of error that the user is making" (David & Armin, 2019). This was also found in our study, as the latter half of the game moved much quicker.

3.3 Distribution Fitting

In order to generate stochastic data by means of a simulation model, the team first had to find service times that best resembled the process. To do this, the research team used Expert Fit and Minitab, entering the data detailing the individual duration of each puzzle step to find a distribution that best fits this data. While the team was able to find strong candidate distributions for each step, tested and unable to be rejected by both the Kolmogorov-Smirnov and Chi Squared tests, it is worth noting that it is hard to reject distributions when there were only at most 10 data points per puzzle. Chi square tests require at least five data points, but the distribution fitting is better applied as the sample size increases (NIST, 2023). These tests for goodness of fit get easier to reject, and thus can be considered to be valid, when considering as much data as possible.

For the teams that did not successfully complete an escape and were not able to record information on for each puzzle, the research team had to consider other forms of distribution fitting, as Expert fit requires an input of at least ten points of data. For those puzzles that did not satisfy this requirement, the research team used Minitab. Despite the small sample size, Minitab was able to determine the best distribution for the remaining puzzles using probability plots (Figure 2) and the goodness of fit test. In order for a sample to fit a particular distribution (with $\alpha = 5\%$), it is desired to have p-values > 0.05, which is the case for the examples shown. As the distributions are created with expert fit and Minitab, it is critical that in the future more data is collected to increase the strength of these approximations. While this data may not be substantial enough in its current state to draw new conclusions, it can be useful to help confirm prior notions to the pace in which a team should progress at, and will lay a foundation for the procedure to retest these general notions when more data is collected.



a) Probability plot for Lion Lock b) Probability plot for Flip Table

Figure 2: Shows probability plots generated for two of the puzzles showing a Logarithmic and Normal Distribution

3.4 Simulation

In order to reflect the process of a group going through an escape room, the team generated simulations of each of the rooms in Simio. Originally, the team had set each puzzle as a separate server within the system, however, there are certain puzzles that can be done at the same time, and the game will not process until all tasks are completed; because these puzzles were not required to be done in a particular order, nor did every team complete them in a similar sequence, these puzzles can theoretically be grouped together, with their service time represented as the time in which the final puzzle within the grouping was finished. The games were also separated into 'phases,' in the same way that the escape rooms are physically oriented by rooms.

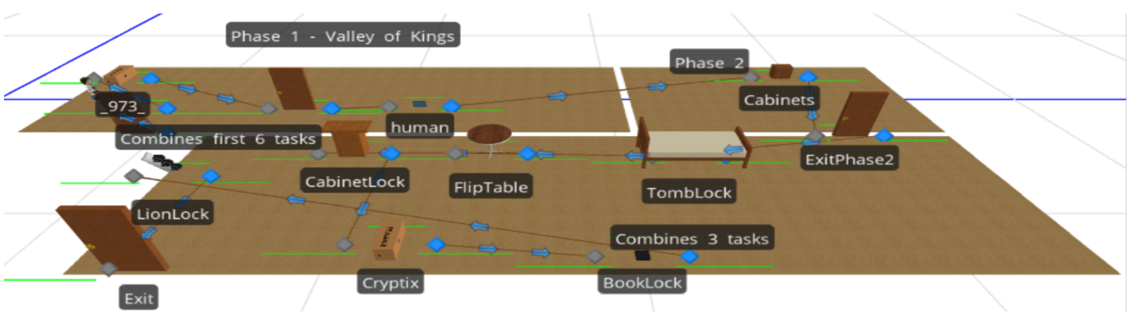


Figure 3: Overhead view of Simio model: Phase 1, 2 and 3 represented.

For the most popular room, it is separated into three phases and has a time limit of 60 minutes and each puzzle, or each group of puzzles, being represented by a server and using connectors to order the steps. From Expert Fit and Minitab, the team was able to obtain a function representing the time it takes to complete each puzzle, which was entered as the service time distribution for the appropriate server. Popular candidate distributions included, but were not limited to, logarithmic, beta, and normal functions. Two models were created for each room. One model follows the original distributions with the values taken from the data collection, and reflects the current state of the room. In this, not all groups are able to escape the room, proven by the simulation running over 60 minutes. This shows that there are current bottlenecks in the system, and that some puzzles currently take too long to make it out of the room. The times from this simulation represent the current process at the escape room center, as data shows that over 50% of teams will not make it out of the room. The second model shows an ideal state of how long each puzzle should take in order for most teams to get close to

ISBN: 97819384964-4-8

getting out. This represents the wish of the client, as they do not want everyone to escape, just far enough that they feel satisfied with the game. To determine the ideal state, only the teams that completed all the puzzles and escaped the room were considered. From these teams, the average time it took to complete each puzzle was calculated; this time is found in the generated ideal pace column in table x. below. After generating the ideal pace, an ideal Simio model was created by adjusting the distributions for each of the puzzles. In the ideal Simio model, the goal is for most teams to be able to escape. So, the research team compared the original pace from Simio with the generated ideal pace to determine what puzzles were creating bottlenecks and not lining up with the ideal data. However, when comparing the data, it was found that in the Simio model, some of the puzzles took less time than in the ideal data. To create the most efficient and ideal model, the research team chose to work with the minimum value for each puzzle. When the generated ideal data was used, the distribution was created using normal distribution. Normal distribution was chosen because the team wanted to account for differences in skill levels among teams with standard deviation. To do this, the generated ideal completion time for the puzzle was used as the mean and the standard deviation previously found using all time studies and not just the ones that were completed, was used. Once the distributions for each puzzle were determined, data using the ideal pace were generated in Simio. This data is shown below in Table 1. The comparison from the original pace to the ideal pace shows that there were some puzzles, for example Door to Nowhere, that took much longer than expected in actual gameplay then predicted. Total average time for all puzzles was 56.33 which is less than 60, showing that on average, most teams would be able to complete the puzzle if kept at this pace. A puzzle that the research team highlighted as needing rework was the Cabinets, as the number of hints necessary and time taken made it a bottleneck to the system. This information can be used by the GM to determine when to give hints to keep teams on track to successfully escape the room.

Table 1. Comparison of paces from original test to ideal pace tested in Simio. Original test represents data that the center has used up until now to pace teams, Simio pace is the times from Simio based on data from the team's time study.

	Original Test	Simio Pace	Generated Ideal Pace	Ideal Pace in Simio
R1 Puzzles	12.46	21:43	14:08	13:39
Sarcophagus Bookcase	11:13	5:49	10:25	9.42
Door to Nowhere	0:40	1:39	1:33	2:46
Skull Box	2:05	2:30	2:08	2:10
Cabinets	7:25	9:34	6:38	6:27
Exit Lock	10:58	7:33	9:34	8:04
Tomb Lock	2:34	4:13	2:43	2:06
Flip Table	5:23	3:21	4:41	4:02
Cabinet Lock	2:40	2:42	2:52	2:42
Cryptix	1:55	3:45	2:44	2:22
Book Lock	1:01	1:21	0:39	0:38
Lion Lock	0:20	1:24	0:23	1:27

4. Discussion

The escape room facility is expanding with the addition of another escape room as well as introducing new volunteer Game Masters. To evaluate and improve the escape room process, the team conducted time studies and generated simulations to monitor the current process and determine ways to improve it. One of the original concerns the client had was the feasibility of each team to escape any of the given rooms. To give everyone a chance to escape, the team evaluated each of the puzzles through time studies and looked at the deviation of how long each one took; the puzzles with a larger deviation were those that were critical in figuring out if the team had a chance to finish or not. The team then highlighted these puzzles as key points that the center should try to make sure they are on pace with to give each team a fair chance of success. This

ISBN: 97819384964-4-8

data will let the Game Master know if a clue should be given to a team that is stuck on a puzzle, based on the likelihood of a group getting out at that moment.

Using the current data taken and evaluated from the time studies, the team generated a current and ideal model for the most popular of the rooms. With this data, the average process time was greater than the 60 minutes allotted for a team to escape. For the ideal model, the team looked at the puzzles that did not line up with the ideal pace determined from the data from teams that escaped. The team then considered the larger deviations from teams that did not perform as well in the room and created distributions that would reflect the ideal pace with these deviations. The team was able to generate a model that had an average completion time of 56.33 minutes showing that the ideal model can be used by the GMs to keep teams on track to complete the room. By recording and organizing the most commonly used hints that were given by the Game Master during the time studies, a new GM will be able to send hints and assist teams on puzzles that teams frequently get stuck on. The simulation data will help to alert a GM when they should be focusing on a game the most, allowing them to better monitor two games at a time from the same control room. Through our efforts, the client will be able to offer an improved customer experience that will result in increased consumer retention and profit margins, boosting the overall product brand. A goal of the engineering work being performed will allow for a lower number of employees to be employed for operation and enable people with little to no experience within an escape room be able to successfully perform as GM, while simultaneously permitting a higher number of hours of operation for consumers to enjoy escape rooms.

5. Future Work

With the constant changing of customers, the escape room center has the opportunity to collect more data past the completion of the project. In the future if another team was to work with the center, a defining initiative could be to complete simulation modeling of all five rooms, and use that data to continue the work being done. This could include more accurate data on how group size, age and other factors affect the probability of a team making it out, and a more in-depth study of how and when hints are sent affects game play. If hints are provided too early or in too high a quantity, customers will have an overall lower amount of enjoyment while playing the game, however not releasing the hints when needed will hinder the customers' ability to complete the game.

6. Conclusion

In conclusion, the team was able to develop and deliver detailed and thoroughly developed models of the most popular room within the escape room center, with the goal of providing the client with data on how to allow the team to get close to finishing a room. The first model that was built signaled that not many groups would not be able to complete the game, as the simulation run time was over 60 minutes. This backed up our findings in the data that over 50% of teams that attempted to escape the room were unable to do so. Building off of this, the second developed model attempted to determine the "ideal" times that each puzzle should take within the room to allow for the room to be escaped more often. These ideal times were taken from comparing the Simio model pace and times recorded from where teams escaped, and using varying distributions depending on which time was faster. Following the developed methodology, the average time to completion was 56:20. This being less than the 60 minutes allocated for the room is reasonable, as it works as a basis to follow before teams face difficulties and when hints are necessary.

With the escape room center having the primary directive of providing to a charity, it was imperative that the team was able to apply these techniques in order to produce a measurable improvement within the operation. Through this improvement, the team believes an increase in generated revenue will be possible and in turn will result in larger contributions to the nonprofit aiding special needs individuals. Additionally, the work done by the team will be beneficial to new employees learning to run and oversee each of the puzzle rooms. The teams' work provides a framework that can guide the decision making of the Game Master, ensuring that any group playing a game will not go without assistance if it is required as well as ensuring maximum enjoyment.

7. References

Birkun, A. et al. (2018, September 15). *Pre-recorded instructional audio vs. dispatchers' conversational assistance in telephone cardiopulmonary resuscitation: A randomized controlled simulation study*. World Journal of Emergency Medicine. Retrieved April 17, 2023, from http://wjem.com.cn/EN/10.5847/wjem.j.1920-8642.2018.03.001

David, D., & Armin, E. (2019). Development of Escape Room Game using VR Technology. *Procedia Computer Science*, *157*(2019), 646-652. https://www.sciencedirect.com/science/article/pii/S1877050919311421

Faber, J., & Fonesca, L. (2014). How sample size influences research outcomes. *Dental Press J Orthod*, *2014*(Jul-Aug), 27-

29. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4296634/

National Institute of Science and Technology (NIST). (n.d.). *Chi-Square Goodness-of-Fit Test*. 1.3.5.15. Chi-square goodness-of-fit test. Retrieved March 20, 2023, from https://www.itl.nist.gov/div898/handbook//eda/section3/eda35f.htm

Winkler, G. M. R. (1993). *Introduction to Robust Statistics and Data Filtering*. Hamilton Technical Services Home Page. Retrieved March 28, 2023, from https://www.wriley.com/ROBSTAT.htm

Veldkamp, A. et al. (2020). Escape education: A systematic review on escape rooms in education. *Educational Research Review*, *31*(2020), 1. https://www.sciencedirect.com/journal/educational-research-review